# Poster: Privacy-Aware Decentralized Multi-Slice Traffic Forecasting

Hnin Pann Phyu, Diala Naboulsi, Razvan Stanica

# Poster: Privacy-Aware Decentralized Multi-Slice Traffic Forecasting

HNIN PANN PHYU, École de Technologie Supérieure (ÉTS), Canada
DIALA NABOULSI, École de Technologie Supérieure (ÉTS), Canada
RAZVAN STANICA, Univ Lyon, INSA Lyon, Inria, CITI, France

In this work, taking the perspective of Mobile Virtual Network Operators (MVNOs), we tackle the multi-slice traffic forecasting problem, while respecting the data privacy of users. To this end, we propose the Federated Proximal Long Short-Term Memory (FPLSTM) framework, which allows MVNOs to train at each base station their local models with their private datasets, without compromising data privacy. Prediction results obtained by evaluating the models on a real-world dataset indicate that the forecast of FPLSTM is as accurate as state-of-the-art solutions while ensuring data privacy as well as computation and communication costs efficiency.

CCS Concepts: • **Networks** → **Network management**.

Additional Key Words and Phrases: network slicing, traffic forecasting, data privacy, machine learning

## 1 INTRODUCTION

Network slicing is designed to leverage the development of future mobile communication services, i.e. enhanced Mobile BroadBand (eMBB), Ultra-Reliable and Low-Latency Communication (URLLC), and massive Machine-Type Communication (mMTC), by satisfying diverse Quality of Service (QoS) requirements [4]. Since these different applications exhibit diverse traffic patterns, achieving an efficient resource allocation, while meeting QoS requirements remains a critical challenge for MVNOs [2]. To this end, traffic forecasting for individual slices can assist MVNOs in the resource reconfiguration process at each base station, thereby enhancing the resources utilization in the Radio Access Network (RAN).

One promising solution to overcome the foregoing challenges is a decentralized collaborative machine learning scheme, also known as Federated Learning (FL) [3]. With this in mind, we propose a privacy-aware FL-based forecasting algorithm that can be used by MVNOs for base stations level traffic predictions. Our solution integrates FL with a well-known time-series forecasting technique, Long Short-Term Memory (LSTM), to enhance the forecasting accuracy of multi-slice network traffic. In the following sections, we lay out the hypothesis of our proposed solution and some initial experimental results.

## 2 SYSTEM MODEL

We consider $\mathcal{B}$ as a set of heterogeneous base stations, managed by the Infrastructure Provider (InP). A base station $b \in \mathcal{B}$ could thus cover a macro-, micro-, pico-, or femto-cell. We use $s \in \mathcal{S}$ to represent a slice instance and $\mathcal{S}$ to represent the set of slice instances. Without loss of generality, we consider each MVNO offers only one service through a slice instance $s$ on a set of base stations $\mathcal{B}_s \subseteq \mathcal{B}$. We consider time is slotted and we use $t$ to refer to one time interval and $T$ to refer to a set of time intervals of interest. We denote as $v_b^s(t)$ the generated traffic volume at base station $b \in \mathcal{B}_s$ of slice instance $s \in \mathcal{S}$ at time $t \in \mathcal{T}$.

In what follows, $X_b^s$ denotes the private dataset, representing historical traffic information, collected at base station $b$ for slice $s$. We denote $X_b^s(t,H) = \{v_b^s(t-H), ..., v_b^s(t-1)\}$ as the historical evolution of the traffic, with $H$ representing an observation window. Accordingly, the associated forecast traffic for a prediction window $W$ is obtained as: $\tilde{X}_b^s(t,W) = J_b^s(X_b^s(t,H))$, where $J_b^s(.)$ is the forecasting model associated to base station $b$ of slice $s$.

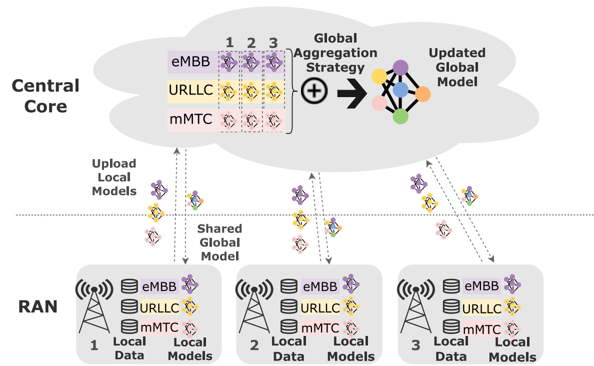


Fig. 1. FL-driven slice traffic forecasting architecture.

### 2.1 Proposed Framework

We design the privacy-aware FL approach as envisioned in Fig. 1. Its operations are conducted at every time time $t \in \mathcal{T}$. Specifically, each local agent (an instance per MVNO and per base station) trains its own local model with its own local dataset. After that, the global agent (managed by the InP) aggregates the uploaded weight matrix from the local models in order to update its global model. Finally, the updated global model is shared with all participating local agents.

Technically, the overall objective can be represented as:

$$min_{w_J}\ GL(w_J) = \sum_{s \in \mathcal{S}} \sum_{b \in \mathcal{B}_s} \rho_b^s\ LL(w_{J_b^s}) \tag{1}$$

where $GL(.)$ and $LL(.)$ are the global loss function and local loss function (i.e. root mean squared error (RMSE) in our case) of global model and local model, respectively. We denote $w_J$ as the weight matrix of the neural network acting as global model $J$ (i.e. LSTM in our case) and $w_{J_b^s}$ is the weight matrix of the local model $J_b^s$, while $\rho_b^s$ is the relative impact of each local agent.

## 3 EXPERIMENTAL RESULTS

We derive our results over 57 base stations of the Orange 4G network located in Poitiers, France. We first compare the performance of our FPLSTM solution with the best known centralised counterpart from the state of the art, Deepcog [1]. Fig. 2 depicts the average RMSE values of FPLSTM and Deepcog for four considered slices, representing Facebook, Youtube, Google and Instagram traffic. As exhibited in this figure, the RMSE values of FPLSTM is indeed slightly higher, but comparable, to those obtained by the centralized Deepcog solution. However, we note that in Deepcog the MVNOs are required to share possibly sensitive and private data. In Fig. 3, we also plot the resulting predicted traffic trend and associated ground truth for the Facebook slice, showing that the performance of the two solutions is really close.

As shown in Fig. 4, the communication cost of FPLSTM is lower, we use a lower fraction of the dataset for training, as no actual data is transferred between nodes. On the one hand, the communication cost of Deepcog increases significantly with the fraction of the dataset used in the training process. Moreover, the resulting RMSE values in the figure show that FPLSTM exhibits better sample efficiency than Deepcog. Fig. 5 depicts the utilization time per CPU, in hours, of FPLSTM and Deepcog for the four types of slices. Here, we observe that CPU utilization of FPLSTM is approximately 2 to 6 times better than Deepcog.

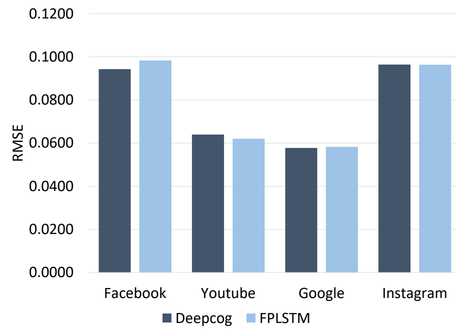


Fig. 2. RMSE values for FPLSTM and Deepcog

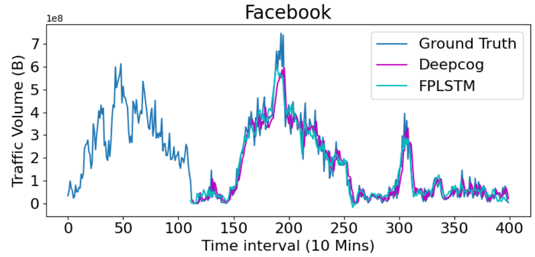


Fig. 3. Forecasting results vs. Ground truth for Facebook traffic using FPLSTM and Deepcog.

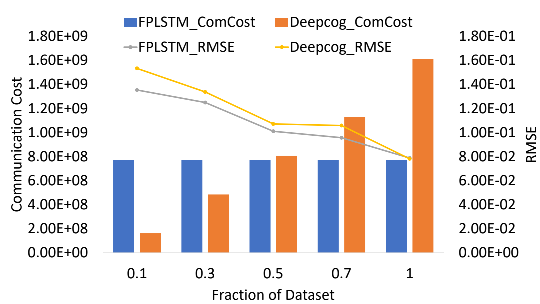


Fig. 4. Comparisons of FPLSTM and Deepcog under different fractions of data: communication cost and RMSE.

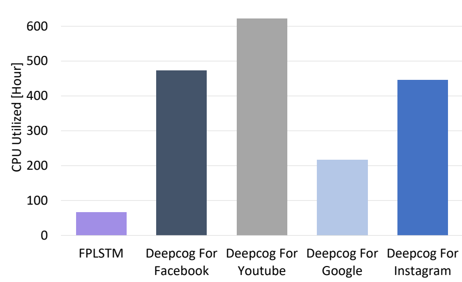


Fig. 5. FPLSTM and Deepcog per CPU utilization time.

## 4 CONCLUSION

In this paper, we address the multi-slice traffic forecasting problem, to facilitate intelligent and anticipatory resource management. According to a series of experiments on a real-world dataset, the performance of FPLSTM is shown to be as accurate as that of a state-of-the-art centralized solution, while ensuring data privacy, and improving communication and computation efficiency.

## ACKNOWLEDGMENTS

This work was supported by the École de Technologie Supérieure (ÉTS), the National Natural Sciences and Engineering Research Council of Canada (NSERC) through research grant RGPIN-2020-06050 and by the ANR CANCAN (ANR-18-CE25-0011) and the ANR MAESTRO5G (ANR- 18-CE25-0012) projects.

## REFERENCES

[1] Dario Bega, Marco Gramaglia, Marco Fiore, Albert Banchs, and Xavier Costa-Perez. 2019. DeepCog: Cognitive Network Management in Sliced 5G Networks with Deep Learning. In *Proc. IEEE Infocom*. Paris, France.

[2] Wanqing Guan, Haijun Zhang, and Victor C.M. Leung. 2020. Slice Reconfiguration based on Demand Prediction with Dueling Deep Reinforcement Learning. In *Proc. IEEE Globecom*. Taipei, Taiwan.

[3] Ahmed Imteaj, Urmish Thakker, Shiqiang Wang, Jian Li, and M. Hadi Amini. 2022. A Survey on Federated Learning for Resource-Constrained IoT Devices. *IEEE Internet of Things Journal* 9, 1 (2022), 1–24.

[4] NGMN. 2016. *Description of Network Slicing Concept*. White Paper. NGMN 5G Project Requirements & Architecture – Work Stream E2E Architecture.